Using Machine Learning to Predict Lung Transplant Graft Failure

Student Researcher: Catherine Michelutti, School of Engineering and Applied Science and The Wharton School, class of 2023
Mentor: Dr. Hersh Sagreiya, Assistant Professor, Perelman School of Medicine, Dept. of Radiology

Funded by the Penn Undergraduate Research Mentorship Program (PURM)

Introduction

- Goal: Create a machine learning model that can most accurately predict lung transplant graft failure or success based on information about a patient's health.
- Primary graft dysfunction (PGD) affects 10-25% of lung transplant patients.
- PGD is a primary cause of post transplant mortality.
- According to a University of Michigan lab, the five-year survival rate of lung transplantation is 50% and the ten-year survival rate drops to 20%.
- Increased accuracy of survival predictions may lead to more efficient lung assignments.

Data/Data Preparation

- The dataset was provided by the United Network for Organ Sharing (UNOS).
- 50 prediction variables were included in the model to predict "gstatus" (graft failure).
- gstatus is a binary variable, 1 = graft failure and 0 = graft success.
- Prediction variables were a combination of numerical and categorical. I transformed categorical variables to binary numerical variables. (Yes/True = 1, No/False = 0).

Statistical Analysis

- The chi-square test only found a significant relationship between gstatus and init_llu_flg, init_blu_flg, end_rlu_flg and end_blu_flg. These variables indicate lung preference at registration and at transplant. "rlu" is right lung, "llu" is left lung and "blu" is both lungs.
- The Wilcoxon Rank-Sum Test found that age is the only numerical variable that is statistically significant. Significant meaning that the distribution of age values is significantly different between graft failure and success.

Predictive Models

- Used a logistic regression, a support vector machine (SVM), a neural network and a random forest classifier.
- After trying SVM models with different kernels, found that a polynomial kernel with degree = 2 performed the best.
- Achieved the best performance with the random forest classifier with 100 trees.
- Used a sequential model with two layers and an output layer for the neural network.
- I combined the random forest classifier, logistic regression and the support vector machine to create an ensemble model.

Discussion and Conclusion

- I initially used principal components analysis (PCA) in hopes of improving the performance of the models. However, I realized that when I removed PCA accuracy and precision increased.
- I switched from using the ROC AUC metric to using the PR AUC metric because PR AUC which improved individual model scores since the data was heavily imbalanced.
- A limitation of the dataset was that there were a lot of missing values. The data started with 178,000 rows. However, after removing all the rows that had NA values within the 50 variables I chose, the number decreased to 941 rows. In the future I should use imputation to fill in empty values so as to have more data points to work with.
- In conclusion, the random forest classifier and the ensemble model were able to predict graft failure most accurately with an accuracy of 0.79.

References

- "Chi-Square Statistic: How to Calculate It / Distribution." *Statistics How To*, 6 July 2020, www.statisticshowto.com/probability-and-statistics/chi-square/.
- Chollet François. Deep Learning with Python. Manning Publications Co., 2018.
- Czakon, Jakub. "F1 Score vs ROC AUC vs Accuracy vs PR AUC: Which Evaluation Metric Should You Choose?" Neptune.ai, 30 July 2020, neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc.
- James, Gareth, et al. *An Introduction to Statistical Learning: with Applications in R*. Springer, 2013..
- Urban, Kylie. "Chronic Rejection in Lung Transplants - the Exploration and Quest to Stop It." *University of Michigan*, 28 Feb. 2017, labblog.uofmhealth.org/lab-report/exploring-cause-of-chronic-lung-transplant-rejection-a-quest-to-stop-it.

Results

Random Forest Classifier:

```
              precision    recall  f1-score   support

         0.0       0.44      0.10      0.17        39
         1.0       0.81      0.97      0.88       150

    accuracy                           0.79       189
   macro avg       0.62      0.53      0.52       189
weighted avg       0.73      0.79      0.73       189

Random Forest PR AUC Score:  0.8051587301587302
```

Logistic Regression:

```
              precision    recall  f1-score   support

         0.0       0.25      0.03      0.05        39
         1.0       0.79      0.98      0.88       150

    accuracy                           0.78       189
   macro avg       0.52      0.50      0.46       189
weighted avg       0.68      0.78      0.71       189

Logistic PR AUC Score:  0.7945757185757186
```

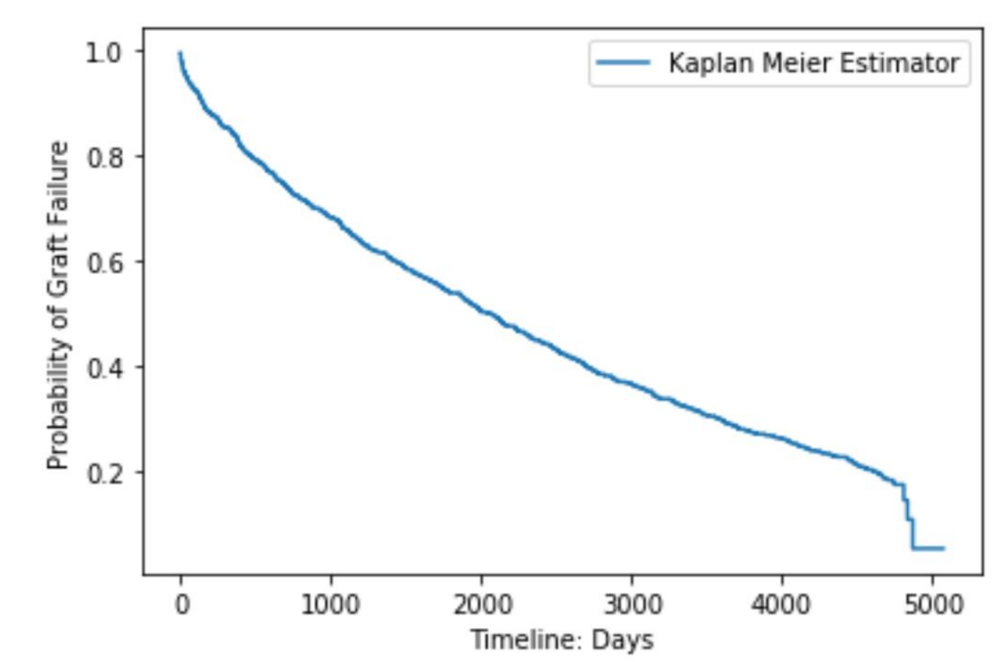
Support Vector Machine:

```
              precision    recall  f1-score   support

         0.0       0.41      0.23      0.30        39
         1.0       0.82      0.91      0.86       150

    accuracy                           0.77       189
   macro avg       0.61      0.57      0.58       189
weighted avg       0.74      0.77      0.75       189

SVM PR AUC Score:  0.8180445458289769
```

Ensemble Model:

```
Accuracy Score:  0.7989417989417989
Ensemble Model PR AUC Score:  0.8042051989878077
```

Kaplan Meier Curve:



Prediction Variable	P Value
hemo_pa_mn_trr	0.226787
hemo_pa_mn_tcr	0.206808
hemo_co_tcr	0.298167
hemo_co_trr	0.411433
hemo_pcw_trr	0.180603
hemo_pcw_tcr	0.137137
init_o2	0.437203
init_creat	0.152153
init_calc_las	0.491277
init_match_las	0.491277
init_bmi_calc	0.077799
tot_serum_album	0.166135
hemo_sys_tcr	0.103632
init_hgt_cm_calc	0.143408
end_bmi_calc	0.063234
age	2.42E-06
end_creat	0.457197
end_calc_las	0.191684
end_match_las	0.199603
gender	0.651835
init_rlu_flg	0.103255
init_llu_flg	0.030787
init_blu_flg	0.000178
ventilator_tcr	0.268762
inotropes_tcr	1
pros_infus_tcr	0.645835
pge_tcr	1
oth_life_sup_tcr	0.417585
ecmo_tcr	0.823323
end_rlu_flg	0.047853
end_llu_flg	0.135957
end_blu_flg	0.000141
ventilator_trr	0.852929
inhaled_no_trr	0.823323
pros_infus_trr	0.664931
pge_trr	1
oth_life_sup_trr	0.280194
cereb_vasc	0.674036
malig_tcr	0.08528
dial_after_list	0.516089
inotrop_vaso_sys_tcr	0.541472
inotrop_vaso_dia_tcr	0.541473
prev_tx	0.394132
prev_tx_any	0.51663
hep_c_anti_don	0.918278
non_hrt_don	0.823323
diab	0.973333